

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

AD

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2000-115772

(P2000-115772A)

(43) 公開日 平成12年4月21日 (2000.4.21)

(51) Int.Cl. ⁷	識別記号	F I	テマコード* (参考)
H 0 4 N 7/24		H 0 4 N 7/13	Z
5/60		5/60	Z

審査請求 未請求 請求項の数11 O L (全 13 頁)

(21) 出願番号 特願平11-244523

(22) 出願日 平成11年8月31日 (1999.8.31)

(31) 優先権主張番号 09/144161

(32) 優先日 平成10年8月31日 (1998.8.31)

(33) 優先権主張国 米国 (US)

(71) 出願人 596092698

ルーセント テクノロジーズ インコーポ
レーテッドアメリカ合衆国, 07974-0636 ニュージ
ャーシイ, マレイ ヒル, マウンテン ア
ヴェニュー 600

(72) 発明者 ヤン ヒュアング

アメリカ合衆国 08817 ニュージャーク
イ, エディソン, ウッドヘイヴン ドライ
ヴ 501

(74) 代理人 100064447

弁理士 岡部 正夫 (外11名)

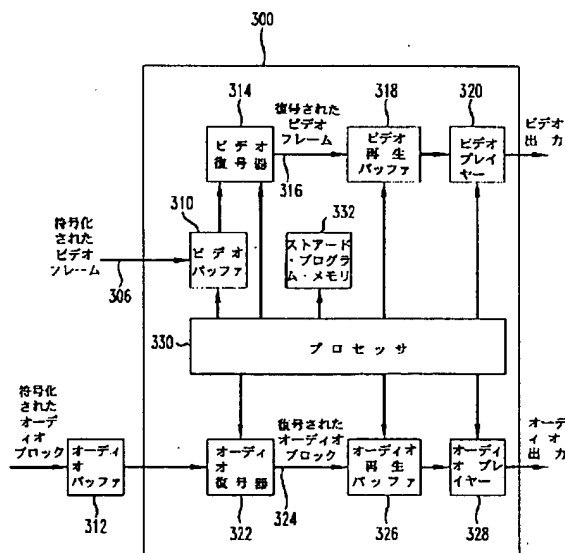
最終頁に続く

(54) 【発明の名称】 マルチメディアの再生システムにおけるデジタルのビデオおよびオーディオの再生の適応型同期化のための方法および装置

(57) 【要約】 (修正有)

【課題】 変動のある再生環境における活動中にユーザ認識の再生性能が改善される、デジタル・ビデオおよびオーディオビデオ及びオーディオ情報の同期化再生に関する。

【解決手段】 ストアード・プログラム・メモリ付きのプロセッサが、オーディオ・ブロックおよびビデオ・フレームの復号および再生を制御し、各ストリームが意図された再生時刻の順序で独立に格納されている。プロセッサはオーディオ・ブロックの復号および再生に対して高い優先度を与え、符号化のビデオ・フレームのバックログの存否を定期的に判定する。ビデオのバックログがある状況下でプロセッサは意図されている再生時刻が最新のバックログのビデオ・フレームのシーケンス内でのフレームを復号し、再生用に復号さるべきバックログのビデオ・フレームだけを復号して、そのバックログを減らす。



【特許請求の範囲】

【請求項1】 再生の乱れを減らすために、符号化されたオーディオおよびビデオの情報の組み合わせられたシーケンスを復号してレンダリングするマルチメディア再生システムにおいて使うための方法であって、該方法は、復号されたオーディオおよびビデオ情報のレンダリングを監視し、それが前記オーディオ情報のレンダリングに関してバックログになっているかどうかを判定するステップ及び、

前記オーディオの情報の復号およびレンダリングに高い優先度を許可するステップ並びに、

前記ビデオ情報が前記オーディオ情報に関してバックログになっている時に、

前記バックログになっているビデオ情報のシーケンスの中で最も最近のビデオ画像をレンダリングするために復号されなければならないバックログになっているビデオ情報の最小の組を決定するステップと、

前記最小のバックログになっているビデオ情報の組だけを復号するステップと、

前記最も最近のビデオ画像だけを前記バックログからレンダリングするステップ及び、

前記最も最近のビデオ画像を再生するためには不要である他のビデオ情報をすべて捨てるステップとによって、前記ビデオのバックログを減らすステップとを特徴とする方法。

【請求項2】 請求項1に記載の方法において、前記ビデオのバックログを減らすステップが、

前記バックログになっている情報が符号化されたビデオ・フレームおよびオーディオ・ブロックのシーケンスを含んでいる場合、各オーディオ・ブロックのレンダリングの直後の時点で前記システムによって格納されているバックログになっているビデオ・フレームの数を表す個数を求めるステップ並びに、

バックログになっているビデオ・フレームの前記個数が規定された数を超えた時に、

前記バックログになっているビデオ・フレームがキー・フレームを含んでいるかどうかを判定するステップと、

前記バックログになっているビデオ・フレームのシーケンスの中で最も最近のビデオ・フレームをレンダリングするために復号されなければならない、前記バックログになっているビデオ・フレームの最小の組を識別するステップと、

前記バックログになっているビデオ・フレームの最小の組を復号するステップ及び、

前記バックログになっているビデオ・フレームの最小の組から前記バックログになっているシーケンスの中の前記最も最近のビデオ・フレームをレンダリングするステップとによって、

前記バックログを消去するステップとを含む方法。

【請求項3】 請求項2に記載の方法において、前記バ

ックログになっているビデオ・フレームのシーケンスが少なくとも1つのビデオ・キー・フレームを含む時、前記バックログになっているビデオ・フレームの最小の組が、前記バックログになっているビデオ・フレームのシーケンスの中で最も最近のキー・フレームおよび、前記バックログになっているフレームのシーケンスの中の前記最も最近のキー・フレームの後に続くすべての他のバックログになっているビデオ・フレームを含む方法。

【請求項4】 請求項2に記載の方法において、前記バックログになっているビデオ・フレームのシーケンスがビデオ・キー・フレームを含んでいない時、前記バックログになっているビデオ・フレームの最小の組がすべてのバックログになっているビデオ・フレームを含む方法。

【請求項5】 請求項2に記載の方法において、前記バックログになっているビデオ・フレームのシーケンスが、再生のためにスケジュールされている時刻が、最新のレンダリングされるビデオ・フレームに対してスケジュールされている再生の時刻より早くなく、レンダリングのためにスケジュールされている次のオーディオ・フレームに対してスケジュールされている再生の時刻より遅くないビデオ・フレームを含む方法。

【請求項6】 請求項5に記載の方法において、前記スケジュールされたオーディオ再生時刻が各オーディオ・ブロックに対するシーケンシャルなインデックス番号、オーディオ・ブロック当たり提供される符号化されたサンプルの数を表している数および、オーディオのサンプリング・レートの関数として決定されるようになって

いる方法。

【請求項7】 請求項5に記載の方法において、前記スケジュールされたビデオ再生時刻が各ビデオ・フレームに対するシーケンシャルなインデックス番号とビデオ・フレーム・レートの関数として決定されるようになって

いる方法。

【請求項8】 ユーザ認識の忠実度が高いデジタルのビデオおよびオーディオの情報をシーケンシャルにレンダリングするためのマルチメディア・システムであって、該システムは、

オーディオ復号および再生のセクションと、ビデオ復号および再生のセクションとを含み、

前記オーディオおよびビデオのシーケンスから選択可能な個数のオーディオ・ブロックを復号し、各オーディオ・ブロックの前記システムによるレンダリングの完了を監視し、

各オーディオ・ブロックのレンダリングの直後の時刻において、前記システムによって格納されているバックログになっているビデオ・フレームの数を決定し、

前記バックログになっているビデオ・フレームがキー・フレームを含んでいるかどうかを判定し、

前記バックログになっているビデオ・フレームの数が指

定された数を超えた時、

前記バックログになっているビデオ・フレームのシーケンスの中で最も最近のビデオ・フレームをレンダリングするために復号されなければならない最小個数の前記バックログになっているビデオ・フレームを識別し、前記最小個数のバックログになっているビデオ・フレームを復号し、前記バックログになっているシーケンスの中の前記最も最近のビデオ・フレームをレンダリングすることによって、

前記バックログを消去するように動作する、ストアード・プログラム・メモリ付きのプロセッサをさらに特徴とするマルチメディア・システム。

【請求項9】 請求項8に記載のマルチメディア・システムにおいて、前記オーディオの復号および再生のセクションが、符号化されたオーディオ・ブロックのためのバッファ、復号器、復号されたブロックのためのバッファおよびオーディオ・プレーヤを含むマルチメディア・システム。

【請求項10】 請求項8に記載のマルチメディア・システムにおいて、前記ビデオの復号および再生のセクションが、符号化されたビデオ・フレームのためのバッファと、復号器と、復号されたフレームに対するバッファと、ビデオ・プレーヤとを含むマルチメディア・システム

【請求項11】 請求項8に記載のマルチメディア・システムにおいて、前記マルチメディア・システムはマルチメディアの再生機能を有するパーソナル・コンピュータ、エンジニアリング・ワークステーションおよびデジタル信号プロセッサから構成されているグループから選択されるようになっているマルチメディア・システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、マルチメディア・システムに関する。特に、本発明は、リソースが限られているか、あるいはリソースが変化している再生環境において動作しながらユーザ認識の再生性能が改善されるような、デジタル・ビデオおよびデジタル・オーディオの情報の同期化された再生に関する。

【0002】

【従来の技術、及び、発明が解決しようとする課題】代表的なマルチメディアの再生の応用において、デジタル・オーディオおよびデジタル・ビデオの信号がソースにおいて生成され、そして混合されたストリームの中、あるいは別々のストリームの中で、インターネットなどの通信ネットワーク上でインターネット・ブラウザおよびマルチメディア再生ツールが装備されているパーソナル・コンピュータなどのインテリジェント再生装置に対して配送される。

【0003】この分野の技術においてよく知られているように、デジタル・オーディオ信号は、通常は、アナログ・オーディオ信号のデジタル・サンプルの順序付けられたシーケンスとして生成される。アナログ信号は固定の間隔で周期的にサンプルされ、その間隔の逆数はオーディオのサンプリング・レートとして知られている。記憶または伝送の目的のために、そのオーディオ・サンプルのシーケンスはオーディオ・ブロックとして知られているオーバーラップしていないセグメントに分割されることがよくあり、各オーディオ・ブロックは所定の数のオーディオ・サンプルを含んでいる。

【0004】デジタル・ビデオ信号は、通常は、デジタル形式における静止画像のシーケンスとして捕捉され、表示される。各ビデオ画像は一般にビデオ・フレームとして知られている。単位時間当たりに捕捉されるビデオ・フレームの数はビデオ・フレーム・レートとして知られている。

【0005】ビデオ・フレームおよびオーディオ・ブロックをマルチメディア再生装置に対して転送する際の効率を良くするために、マルチメディアのソースは一般にこのビデオおよびオーディオの情報を符号化する。この情報はデータ・ネットワーク上で再生装置に対して送信される前に、ソースにおいて、そのパケットに対して送信されるオーディオ・ブロックおよびビデオ・フレームのそれぞれに対する適切な再生シーケンスおよび時刻を示している情報を追加して、データ・パケットの中にカプセル化される（パケット化される）ことが多い。再生装置においては、符号化された情報がパケットから抽出され、そして次に直接復号されて再生されるか、あるいは後で処理するために記憶される。ビデオまたはオーディオの再生装置による復号されたデータの再生は、普通、「レンダリング」と呼ばれている。

【0006】MPEG-I、MPEG-IIおよびH263などの標準規格に反映されているようなほとんどのビデオ符号化技法は、再生装置に対して送信されなければならないビデオ情報の量を圧縮するか、あるいは減らすためにフレーム間符号化を使用する。通常は、各デジタル・ビデオ・フレームの中で符号化されているピクセルのうち、連続フレーム上で変化するのはい部分だけなので、送信されるビデオ情報の量は連続フレーム上で変化するピクセルだけを符号化して送信することによって大幅に削減することができる。そのような符号化技法は、通常は、現在のフレームを完全に復号するために、現在のビデオ・フレームに隣接している1つまたはそれ以上のビデオ・フレームを正常に復号するための再生装置を必要とする。隣接しているビデオ・フレームの復号が非効率的であるか、あるいは失われていた場合、ほとんどのビデオ符号化技法は、ビデオ・ストリームの中に「キー」フレームを定期的に挿入することも行い、そのキー・フレームは隣接しているフレームとは独立に復号

される十分な情報を含んでいて、必要な時に圧縮されたビデオの復号・プロセスを再スタートするために使うことができる。

【0007】ビデオ・フレームはオーディオ・ブロックより大幅に多くのデータを必要とするので、マルチメディアのデータ・ストリームは一般に多数のビデオ・パケットの間に少数のオーディオ・パケットをランダムに散在させて含んでいる。これらの散在しているオーディオ・パケットが再生装置に到着した時、ビデオ・フレームとオーディオ・ブロックの意図された再生シーケンスにおける整合が取れない可能性がある。到着時にビデオおよびオーディオのパケットを再同期化するためにいくつかの方式が適用されている（たとえば、ローズノー（Rosenau）他に対して発行された米国特許第5,598,352号（以下、ローズノーの'352特許と呼ぶ）を参照されたい）。

【0008】再生装置に到着した時にマルチメディアのデータ・ストリーム要素の再同期化が行われる場合でも、エンド・ユーザはマルチメディアの再生性能が不十分であることに気付く可能性がある。再生が始まる前に再生装置においてマルチメディアのデータを完全にストリーム化するか、さもなければシーケンスに従って格納しておくことができるが、装置の処理パワーが小さいか、他にかなりの、競合しているタスクのオーバーヘッドがある場合、マルチメディアの再生を一時的に、あるいは長く継続して中断させることになるCPUのバックログによって多かれ少なかれ影響される可能性がある。より安価で他の多くの同時アプリケーションにサービスすることができる再生装置に対してユーザがメーカーに期待するにつれて、再生装置の性能がマルチメディアのアプリケーションの潜在的な性能を制限し続けることになる。

【0009】

【課題を解決するための手段】マルチメディア再生アプリケーションのエンド・ユーザ認識性能は、処理パワーが制限されているか、あるいは大きく変化するマルチメディア再生装置における同期化されたデジタル・ビデオおよびデジタル・オーディオの情報のレンダリングのための新しい発明によって大幅に改善される。通常、ユーザはビデオ再生ストリームの場合よりもオーディオ再生ストリームにおける中断がより気になるので、マルチメディア再生装置はオーディオ・ブロックの復号およびレンダリングに対して最高の優先度を割り当てる。再生装置の処理リソースが制限されている状況下において、この優先順位付けの方式によって、ビデオの復号およびレンダリングがオーディオの復号およびレンダリングより遅れる（あるいは、関連する「バックログとなる」）ことになる可能性がある。ビデオのレンダリングのバックログが直接監視され、そしてそれが大きい場合、バックログになっているビデオ・フレームのシーケ

ンスの中の最近の（または最新の）ビデオ・フレームだけの限定されたレンダリングを実行することによってバックログが削減される。これはそのバックログになっているビデオ・フレームのうち、最も新しいものをレンダリングするために、復号される必要があるバックログのシーケンスの中のビデオ・フレームだけを復号することによって行われる。バックログになっているフレームのうち最も新しいものを復号するために必要とされないバックログのビデオ・フレームは実質的に捨てられる。

【0010】バックログを知るための再生装置の処理リソースの高度な監視の必要性が回避され、そしてオーディオおよびビデオの再生ストリームに対するユーザが気付く中断が最小化されるので有利である。本発明を採用することによって、マルチメディアのアプリケーションは処理パワーが変化する各種の再生装置において効率的にレンダリングを行うことができる。さらに、与えられた再生装置において、マルチメディアのアプリケーションは他の競合しているタスクによって生じた処理リソースの一時的な損失からより良好に回復することができる。

【0011】

【発明の実施の形態】本発明の以下の特定の例示的实施例の記述と共に添付図面を参照することにより本発明のより完全なる理解が達せられるであろう。首尾一貫性を保つため、そして理解し易くするために、類似しているか、あるいは等価な図のそれぞれの要素は下位の二桁の位置の識別番号が同じになっている（たとえば、図1のビデオ再生バッファ118は図3Aのビデオ再生バッファ318と等価である）。

【0012】以前に注記されたように、従来の技術のマルチメディア再生システムは、通常は、ビデオおよびオーディオのデータの混在するストリームを含んでいる圧縮されて符号化されたデジタル信号を受信する。これらのシステムはこの混合されたデータ・ストリームに関して動作し、そのビデオ・フレームとオーディオ・フレームとを分離して復号し、そして次にこれらのフレームおよびブロックを適切なシーケンスで、そして適切な時刻においてレンダリングする。結果として、ビデオ画像とオーディオの出力は密接に同期して発生され、コヒーレントなマルチメディアの経験をシステム・ユーザに対して提供する。

【0013】そのようなシステムのための基本コンポーネントが図1に示されている（たとえば、ローズノーの'352特許を参照されたい）。このマルチメディア再生システム100においては、符号化されたデジタルのオーディオ/ビデオのデータ・ストリーム102がシステム・復号器104に入力され、システム・復号器104はそのストリームを符号化されたビデオのデータ・ストリーム106と符号化されたオーディオのデータ・ストリーム108とに分割する。

【0014】ビデオ・バッファ110はビデオのデータ・ストリーム106を受け取り、そしてオーディオ・バッファ112はオーディオのデータ・ストリーム108を受け取る。ビデオ・復号器114はビデオ・バッファ110からビデオ・データを読み出し、このデータを復号し、そのデータをビデオ・フレーム116にアセンブルし、そして復号された各ビデオ・フレームをビデオ再生バッファ118の中に入れて、ビデオ・プレーヤ120による再生（レンダリング）を待つ。同様に、オーディオ・復号器122はオーディオ・バッファ112からオーディオ・データを読み出し、このデータを復号し、その復号されたデータをオーディオ・ブロック124の中にアセンブルし、そして復号された各オーディオ・ブロックをオーディオ再生バッファ126の中に入れて、オーディオ・プレーヤ128によるレンダリングを待つ。

【0015】従来の技術において各種のマルチメディアのパケット化および符号化の方式が存在する。図2BはそのデータをMPEG-Iの標準に従ってパケットの中にカプセル化する方法を示している（たとえば、MPEG-Iに対するISO/IEC 11172-1:1993（E）の使用を組み込んでいるローズノーの'352特許を参照されたい）。本発明はMPEG-IIおよびH.263などの各種の異なるマルチメディア標準をサポートするために使えることを理解することができる。

【0016】図2において、データ・ストリーム202はISO層およびバック層を組み込んでいる。ISO層においては、終了コード204が配送されるまで、一連のデータ・パッケージ203が通信される。各パッケージはバック開始コード206およびパケット・データ208を含む。バック層においては、各パッケージ202はバック開始コード206、システム・クロック基準（SCR）210、システム・ヘッダ212およびデータのバケット214を含む。これらのバケット214はビデオ・バケットおよびオーディオ・バケットの両方を含み、ビデオ・バケットはそのデータの性質のために、遙かに数が多い。システム・クロック基準210はシステム・タイムを指示し、そしてオーディオのバケットがビデオのバケットの間に散在している頻度を示しているマルチクロックのレートの基準を提供する。

【0017】図2に示されているように、ビデオおよびオーディオの各バケット214は開始コード・プレフィックス216、ヘッダ・データ218およびペイロード・データ220を含む。各種の間隔においてヘッダ・データ218の中にタイム・スタンプ情報222が含まれており、それはペイロード220がビデオ情報を含んでいる場合はビデオ表示のタイム・スタンプ（VPTS）を表し、あるいはペイロード220がオーディオ情報を含んでいる場合はオーディオ表示のタイム・スタンプ（APTS）を表すことができる。MPEG-Iの標準

に従って、VPTSおよびAPTSが、関連付けられているビデオ・フレームおよびオーディオ・ブロックのそれぞれに対する適切な再生時刻を示すために提供される。

【0018】MPEG-IおよびMPEG-IIおよびH.263などの標準は、ビデオおよびオーディオの信号の両方についてのデジタル符号化/復号化および圧縮/伸長を用意している。効率のために、ほとんどのビデオ符号化技法はフレーム間符号化を使ってネットワーク上で再生装置に対して送信されなければならないビデオ情報の量を圧縮または削減する。たとえば、3種類のフレームの混合されたものがMPEG-I標準において使われている（たとえば、ローズノーの'352特許を参照されたい）。「キー」フレーム（イントラ・フレームまたはIタイプ・フレームとも呼ばれる）がそのフレーム自身についての情報だけを使って符号化される。予測またはPタイプのフレームは過去の基準フレームから予測モデルを使って符号化されたデータを含む。双方向またはBタイプのフレームは、過去のおよび/または未来の基準フレームから予測モデルを使って符号化されたデータを含む。予測モデルを使うことによって、データはPタイプおよびBタイプのフレームの両方において大幅に圧縮することができ、ビデオ・データ伝送の総合速度が改善される。

【0019】再生時のオーディオ・ブロックとビデオ・フレームとの同期化は再生性能のユーザ認識のために重要である。オーディオとビデオとの間の再生シーケンスの不整合が大きくなると、再生が不鮮明になる可能性がある。与えられた時間間隔においてビデオ再生シーケンスをサポートするために必要なデータの量は、オーディオの再生のために必要なデータの量より本来的に大きい。結果として、送信されるマルチメディアのデータ・ストリームの中の符号化されたビデオ・バケットの数はオーディオ・バケットの数より遙かに多い。図1の到来するマルチメディアのデータ・ストリーム106におけるビデオ・バケットの間のオーディオ・バケットがまばらになり過ぎている場合、ビデオ・フレーム116とオーディオ・ブロック124との不整合が生じる可能性がある。従来の技術のマルチメディア再生システムにおいては、VPTSおよびAPTSがビデオとオーディオの再生を同期化するためにSCRと組み合わせてよく使われる。たとえば、いくつかの再生システムにおいては、SCRがマルチメディアの再生装置のクロックを同期化するために使われ、それは次にビデオ・復号器114の中のVPTSおよびオーディオ・復号器122の中のAPTSと比較され、その2つの信号の同期が外れている（すなわち、1つのデータ・ストリームが必要な再生時刻に関して他より遅れてる）度合いが求められる。あるしきい値レベルにおいて、再生装置によってその遅れているデータ・ストリームの先頭にジャンプしてその信号

を再同期化するためのアクションが取られる（たとえば、ローズノーの352特許を参照されたい）。

【0020】この再同期化のプロセスによって、到来するデータ・ストリームの中の混乱が補正されるが、それはマルチメディア再生システムに対する処理リソースの利用可能性の変動のために生じ得る追加の振動には対処しない。パワーの限られたプロセッサまたは他の大きな、そして競合しているタスクのオーバーヘッドによって再生のバックログが生成される可能性があり、かなりの数のビデオ・フレームおよびオーディオ・ブロックがバッファの中に格納されたままになる可能性がある。そのようなバックログが発生すると、エンド・ユーザはそのマルチメディア再生中に瞬時的な、そしてより長く続く中断を経験する可能性がある。

【0021】本発明においては、マルチメディアのアプリケーションのエンド・ユーザが認識する性能が、処理パワーが制限されているか、あるいは大きく変動するマルチメディア再生装置におけるデジタルのビデオおよびデジタルのオーディオのデータ・ストリームをレンダリングするための新しい方法によって大幅に改善される。ユーザはビデオ再生ストリームよりオーディオ再生ストリームの中における中断の方がより気になるので、オーディオ・ブロックが復号およびレンダリングにおいて優先される。ビデオの再生のバックログが存在するかどうかを判定するための、単純でしかも信頼性のある方法が使われ、そしてバックログが発生している場合、そのバックログを減らし、オーディオおよびビデオの再生ストリームを同期化するための効率の良い方式が提供される。この方法の実施形態が図3Aに示されている。

【0022】この実施形態においては、マルチメディア再生システムが代表的なマルチメディアPCであるとみなされる。図3に示されているように、プロセッサ330およびストアード・プログラム・メモリ332が、たとえば、オーディオ・復号器322、オーディオ・プレーヤ328、ビデオ・復号器314およびビデオ・プレーヤ320などの各種のシステム構成要素によって共有されている。共有されているリソースとして、プロセッサ330は個別の要素に対して専用されておらず、あるいは任意の特定の時刻において任意の要素に利用できることが保証されてもいない。したがって、構成要素は共有されているリソースに対して競合し、再生の中断が生じる可能性がある。

【0023】従来の技術のマルチメディア再生システムにおいては、そのような競合は実行されるべき復号およびレンダリングのタスクのそれぞれに対して優先度を割り当て、そして次に、これらの優先度に従ってその共有リソースの使用を割り当てることによって対策されてきた。しかし、この方法は構成要素の1つにおいて予期しない処理バックログが発生した場合の調整のための用意はない。本発明においては、図3Aのプロセッサ330

は、ストアード・プログラム・メモリ332によって指示される場合に、優先度を設定し、バックログを減らすために、復号およびレンダリングのタスクについてのいくつかの追加の制御を引き受ける。

【0024】より詳しく言えば、プロセッサ330は、メモリ332によって指示される場合に、オーディオ復号器322に、オーディオ・バッファ312から選択された個数Nの符号化されたオーディオ・ブロック324を呼び出し、これらのN個のオーディオ・ブロック324を復号し、そしてその後、それらをオーディオ再生バッファ326の中に入れるように指令する。オーディオ再生の中断を減らすために、この選択された数Nは少なくともそのシステムに関連付けられているオーディオ再生バッファリングの段数と同程度の大きさで、通常は、2から4までの整数であるべきであることが分かっている。復号の後、プロセッサ330はこれらの復号されたオーディオ・ブロックを再生するようにオーディオ・プレーヤ328に指令する。さらに、プロセッサ330は再生バッファ326の中に残っている復号されたブロックの数を監視し、必要に応じて再生バッファ326を再充填するために十分な数のオーディオ・ブロックを復号するよう、復号器322に指令する。

【0025】各オーディオ・ブロックが再生された後、プロセッサ330はビデオ・バッファ310を調べてバックログになっている符号化されたビデオ・フレーム306の数を数える。バックログになっている符号化されたフレームの数が指定のしきい値を超えた場合、プロセッサは次にバッファ310の中に滞留しているフレームのどれがビデオ・キー・フレームであるかを判定する。

【0026】バッファ310がバックログになっているビデオ・キー・フレームを含んでいた場合、プロセッサ330は、ビデオ復号器314に、そのバックログになっているシーケンスの中の最高のシーケンス番号の、したがって、そのシーケンスの中で最も最近のキー・フレーム（以下、最新のキー・フレームと呼ばれる）を復号し、同様にこのキー・フレームの後にある、したがって、バックログになっているシーケンスの中の後の方にある、バッファ310の中のすべての他のフレームを同様に復号するように指令する。次に、プロセッサ330はビデオ再生バッファ318の中のシーケンス番号が最高である復号されたフレームだけを入れるようにビデオ・復号器314に指令する。したがって、このフレームがその復号されたフレーム・シーケンスの中の最近のものであり、したがって、これ以降では最近の復号されたフレームと呼ばれる。最後にプロセッサ330はこの最近の復号されたフレームを再生するようにビデオ・プレーヤ320に指令する。結果として、最近の復号されたフレームを復号するのに必要なバックログになっているフレームだけが復号され、そしてその最近の復号されたフレームだけが再生される。他のすべてのバックログに

なっているフレームは実効的には捨てられる。

【0027】バッファ310の中にキー・フレームが含まれていなかった場合、プロセッサ330は、ビデオ復号器314に、ビデオ・バッファ310の中のすべてのバックログになっているフレームを復号し、そしてそのバックログになっているシーケンスの中で最も最近の復号されたフレームを再生バッファ318の中に入れるように指令する。次に、バッファ310の中のバックログになっている適切なビデオ・フレームが復号されると、プロセッサ330は最近の復号されたフレームをバッファ318からレンダリングのために呼び出すよう、ビデオ・プレーヤ320に指令する。

【0028】システム300によって操作されるデータ・ストリームのための基本フォーマットが図3Bに示されている。このフォーマットは本発明による1つの例を提供する。各種の他のデータ・フォーマットも可能であり、本発明の範囲内に含まれる。

【0029】データ・ストリーム301はデータ・パケット311を含み、そして各パケットはそのパケットがオーディオのデータまたはビデオのデータのいずれを運んでいるかを示す識別子313および、そのオーディオおよびビデオのパケットがそれぞれパケット化されて記

$$t_{A_i} = i * (\text{オーディオ・ブロック当たりのデータ・サンプル数} / \text{オーディオ・サンプリング・レート})$$

ここで、ブロック当たりのデータ・サンプルの数309およびオーディオ・サンプリング・レート307（図3Bの）がデータ・ストリーム301のストリーム・ヘッダ303の中に設定されている。サンプリング・レートおよびブロック当たりのサンプル数は選択された標準によって変化する。たとえば、MPEG-Iでの「CD品質」のオーディオ伝送は44.1キロヘルツ（kHz）のレートでサンプルされることが可能であり、そして1つのブロックの中に1024個のサンプルを含む。シーケンス番号またはタイム・スタンプのいずれかが関連付けられた符号化されたフレームまたはブロックと一緒に、それぞれ図3Aのバッファ318、326の中に格納される。

【0032】図3Cのビデオ・フレーム356を表している符号化されたデータが到着し、そして図3Aのビデオ・バッファ310の中にシーケンスの順序で入れられる。ふたたび、シーケンスは図3Bに記述されているシーケンス番号315を使うことによって、あるいは何らかの代替手段によって設定することができる。図3Cに示されているように、ビデオのタイムライン358は各ビデオ・フレーム V_i に対するタイム・スタンプ t_{V_i} を伴って設定される。ここで、タイム・スタンプはビデオ・フレーム V_i の再生が開始される所望の時刻を表し、そしてそれは以下の式によって計算される。

【0033】

$$t_{V_i} = j / (\text{ビデオ・フレーム・レート}) \quad (2)$$

*憶されるべき順序を設定するシーケンス番号315を含む。さらに、各パケット311はパケット・ヘッダ317およびデータ・ペイロード319を含む。データ・ストリーム301の全体がストリーム・ヘッダ303から開始される。ストリーム・ヘッダ303はビデオ・フレーム・レート305、オーディオ・サンプリング・レート307およびオーディオ・ブロック・サイズ309を含んでいる。

【0030】図3Cはバックログの状態がどのように判定されるかを示している。オーディオ・ブロック352を表している符号化されたデータが到着し、図3Cのオーディオ・バッファ312の中にシーケンスの順序で入れられる。シーケンスは図3Bに記述されているシーケンス番号315を使うことによって、あるいは何らかの代替りの手段によって設定することができる。図3Cに示されているように、オーディオのタイムライン354が各オーディオ・ブロックのシーケンス番号 i に対してタイム・スタンプ t_{A_i} で設定される。ここでタイム・スタンプは、オーディオ・ブロック A_i が開始される再生のための所望の時刻を表し、そしてそれは以下の式によって計算される。

【0031】

(1)

ここで、図3Bのビデオ・フレーム・レート305はデータ・ストリーム301のストリーム・ヘッダ303の中に含まれている。ビデオ・フレーム・レートは選択された標準および再生装置の機能に従って変化する。たとえば、高品質のビデオ・フレーム・レートは30フレーム/秒（fps）で設定されることが可能である。このフレーム・レートは多くの現在のデスクトップ・マルチメディア・システムの能力を超える可能性があるので、10～15fpsのフレーム・レートがより通常である。

【0034】図3Cを続けて参照すると、現在のオーディオ再生ポインタ360は図3Aのオーディオ・バッファ312からの復号されるべき次のオーディオ・ブロックを識別し、そして現在のビデオ再生ポインタ362は図3Aのビデオ・バッファ310からの復号されるべき次のビデオ・フレーム V_i を識別する。タイム・スタンプ t_{V_i} は V_i に対して決定され、そしてタイム・スタンプ t_{A_i} は次の復号されるべきオーディオ・フレーム A_i に対して決定される。図3Cのビデオのバックログ364は以下のように構成される。

【0035】

$$\text{ビデオ・フレーム } V_i \quad (3)$$

ただし、

$$b \in (\text{シーケンス } p, q)$$

$$t_{V_i} \leq t_{V_b} < t_{A_b}$$

【0036】図3Cに示されているように、ビデオのバ

ックログ364はV₀から始まってV_nに終わっているビデオ・フレームのシーケンスから構成されている。ビデオのバックログは図6に示されていて、以下にさらに記述されるビデオの復号およびレンダリングのプロセスを適応させるために使われる。

【0037】ユーザはビデオ再生ストリームの場合よりもオーディオ再生ストリームにおける中断の方がより気になるので、オーディオの復号および再生のプロセスにはビデオの復号および再生のプロセスよりも高い優先度が与えられる。図4はオーディオ・データの復号・プロセスについてのフローチャートを示している。

【0038】図4のステップ402において、図3Aのオーディオ・復号器322はN個のオーディオ・ブロックをオーディオ・バッファ312から呼び出して復号する。図4のステップ404において、図3Aのオーディオ・復号器322はこれらのN個の復号されたブロックをオーディオ再生バッファ326の中に入れ、そして図4のステップ406において、オーディオ・イベントの到着を待つ。オーディオ・イベントはオーディオおよびビデオの復号および再生のプロセスに対する基本クロックまたはトリガのメカニズムとして働く。

【0039】呼び出されるオーディオ・ブロックの数Nは、通常は、そのマルチメディア・システムによって設定されるシステム・パラメータであり、オーディオの再生の中断をほとんどなくするために十分に大きい必要がある。再生装置のためにデスクトップPCを採用している多くのマルチメディア・アプリケーションに対しては、2〜4ブロックの値の整数Nが有効であることが分かっている。

【0040】オーディオの再生プロセスを示しているフローチャートが図5に与えられ、そしてオーディオ・イベントを発生するために必要なステップを含む。図5のステップ502において、図3Aのオーディオ・プレーヤ328は復号されたオーディオ・ブロックをオーディオ再生バッファ326から呼び出し、そしてこのブロックをオーディオの出力装置を介してステップ504において再生する。図5の判定ステップ506において、図3Aのオーディオ・プレーヤ328はオーディオ・ブロックの再生の完了を監視し、そして完了時にオーディオ・イベントをステップ508において発生する。このオーディオ・イベントを発生した後、プロセッサ330はオーディオ再生ポインタをインクリメントし、図5のステップ502へ戻って次の復号されたオーディオ・ブロックを呼び出す。関連付けられたバッファの設計に合わせて、オーディオ・ブロックはそれが復号されてレンダリングされた後、オーディオ・バッファおよびオーディオ再生バッファのそれぞれから実効的に捨てられる。

【0041】オーディオ再生プロセスによって発生されるオーディオ・イベントは、図3Aのオーディオ・復号器322によって、図4に示されている判定のステップ

406において検出される。ステップ408において、図3Aのオーディオ・復号器322は、オーディオの再生バッファ326がふたたびN個の復号されたオーディオ・ブロックを含むまで、追加の数の符号化されたオーディオ・ブロックをオーディオ・バッファ312から呼び出して復号する。次に、オーディオ・復号器322は図4の判定のステップ406へ戻って次のオーディオ・イベントの到着を待ち、その後、ふたたびステップ408へ移って図3Aのオーディオ再生バッファ326をふたたび満たす。

【0042】ビデオの復号および再生のプロセスを示しているフローチャートが図6に示されている。ステップ602において、図3Aのビデオ・復号器が第1のビデオ・フレームを復号し、そしてこのフレームをビデオ・プレーヤ320によって再生するためにビデオ再生バッファ318の中に格納する。図6のステップ604において、図3Aのビデオ・復号器314は次にオーディオ・イベントの到着を待つ。オーディオ・イベントの到着を検出した後、ビデオ・復号器314は図6のステップ606へ移って図3の現在のオーディオ再生ポインタ310をインクリメントし、ビデオ・バックログ314を判定する。図6のステップ608において、図3Aのビデオ・復号器314はそのビデオ・バックログのサイズに従って2つのアクションのうちの1つを取る。

【0043】そのビデオ・バックログが整数Mより少なかった場合、ビデオ・復号器314は図6のステップ604へ戻り、次のオーディオ・イベントの到着を待つ。ビデオ・バックログがMに等しいか、あるいはMより大きかった場合、ステップ610においてビデオ・復号器314はバックログになっているフレームをサーチしてビデオのキー・フレームがあるかどうかを判定する。

【0044】ビデオのキー・フレームが見つかった場合、図3Aのビデオ・復号器314は図6のステップ612において最も最近のバックログになっているキー・フレームV_kを選択し（タイム・スタンプt_kに基づいて）、最も最近のバックログになっているビデオ・フレームV_nを復号し終わるまで、そのシーケンスの中で順方向に復号を開始し、そしてこの復号された最も最近のバックログになっているフレームV_nをビデオ再生バッファ318の中に入れる。ビデオのキー・フレームが見つからなかった場合、図1のビデオ・復号器314は図6のステップ614においてバックログV₀〜V_nの中のスべてのビデオ・フレームを復号し、そしてバックログになっていた最も最近のビデオ・フレームV_nをビデオ再生バッファ318の中に入れる。いずれの場合においても、最も最近のバックログになっているビデオ・フレームV_nを構成するのに必要なビデオ情報だけが復号される。これらの必要なバックログになっているフレームが復号された後、図3Aのビデオ・バッファ310の設計に従って、バックログになっているすべてのフレーム

10

20

30

40

50

V₀～V₄がバッファ310から実効的に捨てられる。

【0045】図6のステップ616において、図3Aのビデオ・プレーヤ320は最も最近の復号されたフレームV₀をビデオ再生バッファ318から呼び出し、このフレームを再生し、そして図3Bのビデオ再生ポインタ312をインクリメントする。再生バッファ318の設計に従って、復号されたフレームV₀はビデオ・プレーヤ320によってレンダリングされた後、再生バッファ318から実効的に捨てられる。

【0046】ビデオのバックログに対するしきい値レベルを設定する整数Mは、ビデオ再生における大きな不連続性が最小化されるように設定される必要がある。代表的なPCベースの環境においては、Mは1の整数値に設定するのが最も効率的であると信じられる。

【0047】この分野の技術に熟達した人にとっては明らかであるように、図4～図6に示されている本発明に従って、マルチメディア再生装置の中のオーディオの復号および再生のプロセスは、ビデオの復号および再生のプロセスより高い優先度が与えられている。オーディオの再生は本質的に中断されない優先度が与えられているので、各オーディオ再生サイクルは比較的固定された、そして予測可能な期間において完了する。したがって、各サイクルの完了がオーディオおよびビデオの再生および復号のプロセスの両方を制御する「オーディオ・イベント」をトリガするためのクロックとして使われる。ビデオの復号のバックログが検出されると、図3Aのビデオ・復号器314がそのバックログのシーケンスの中の最も最近のビデオ・フレームV₀を復号して表示するためには必要でないすべてのビデオ・フレームをスキップすることができるようにすることによって、そしてビデオ・プレーヤ320がこの最も最近のビデオ・フレームV₀だけを再生できるようにすることによって、オーディオのブロックとビデオのフレームとの間の同期化が再確立される。オーディオ・ブロックの復号およびレンダリングを優先させることによって、オーディオの再生においてユーザが経験する遅延が最小化される。オーディオ再生における遅延は、ビデオ再生シーケンスの中で小さな遅延または省略が発生する場合より、システム・ユーザにとって遙かに気になるものであることが分かっている。

【0048】この上記の方法の実施形態は、本発明の多くの代替実施形態のうちの1つに過ぎないことはこの分

野の技術に熟達した人にとっては前記の説明から明らかである。したがって、この説明は例を示すように作られているだけであり、本発明を実施する最善のモードをこの分野の技術に熟達した人に説明する目的のためのものである。各種の他の代替案が本発明の内容から逸脱することなしに、この分野の技術に熟達した人によって考えられる。たとえば、本発明は、オーディオのブロックおよびビデオのフレームをそれぞれ表している符号化されて復号されたデータの両方を、それぞれが格納することができる単独のオーディオ・バッファおよび単独のビデオ・バッファを使って実装することができる。追加の例として、ビデオ・フレームのバックログに対して選択される値(M)、呼び出されるオーディオ・ブロックの数(N)および、オーディオ・イベント・サイクル当たりの再生されるオーディオ・ブロックの数(1)をここで推奨された値から変更することができる。また、この分野の技術に熟達した人であれば、本発明によって開示されている方法を、オーディオの復号およびレンダリングではなく、ビデオの復号およびレンダリングを優先させるように容易に適応させることができる。

【図面の簡単な説明】

【図1】従来の技術のマルチメディア再生システムの略ブロック図を示す。

【図2】高レベルの構成およびMPEG-I標準に適合している、代表的な従来の技術の符号化されたマルチメディアのデータ・ストリームのフォーマットを示す。

【図3A】本発明の1つの実施形態を示す。

【図3B】本発明の実施形態によって使われているバッケージ化されたオーディオおよびビデオの高い水準の編成の図を提供している。

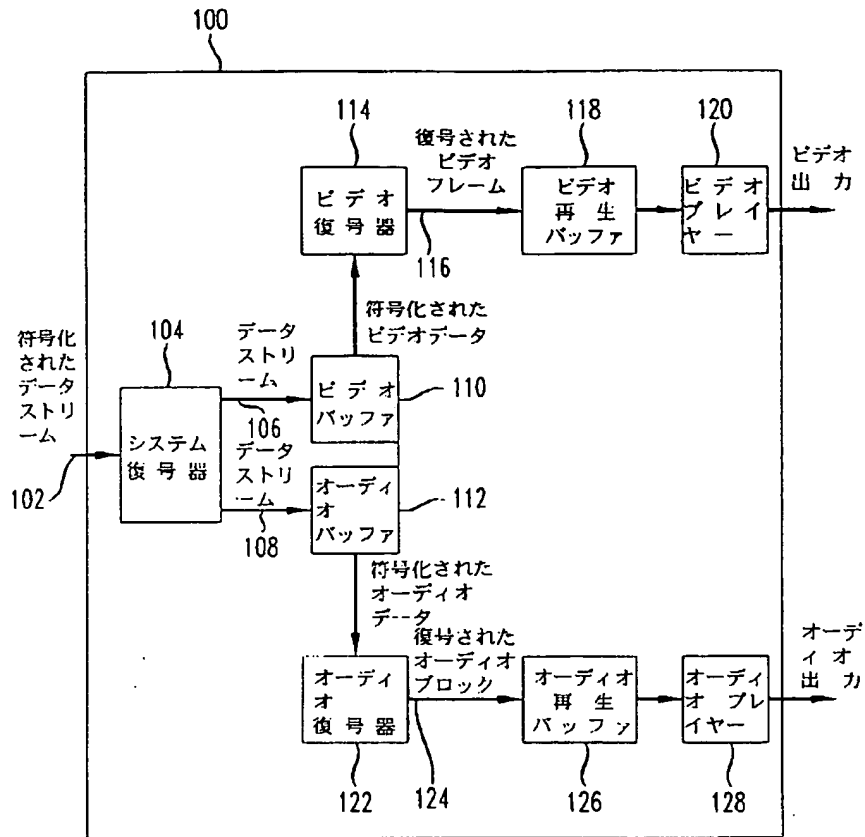
【図3C】本発明によって動作するオーディオ再生とビデオ再生とのストリームの間の時間的關係を図で示す。

【図4】オーディオブロックの復号に使用される本発明のプロセスを示すフローチャートである。

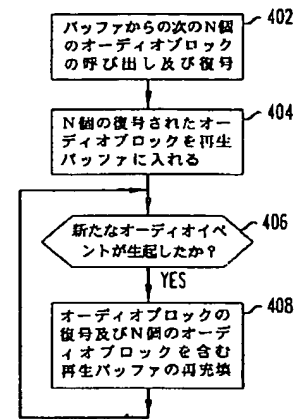
【図5】オーディオブロックのレンダリングに使用される本発明のプロセスを示すフローチャートである。

【図6】ビデオフレームのレンダリングがオーディオブロックのレンダリングに対して大きな遅延を有する場合に、ビデオフレームの復号及びレンダリング並びにビデオ及びオーディオ再生ポインタの再同期化に使用される本発明のプロセスを示すフローチャートである。

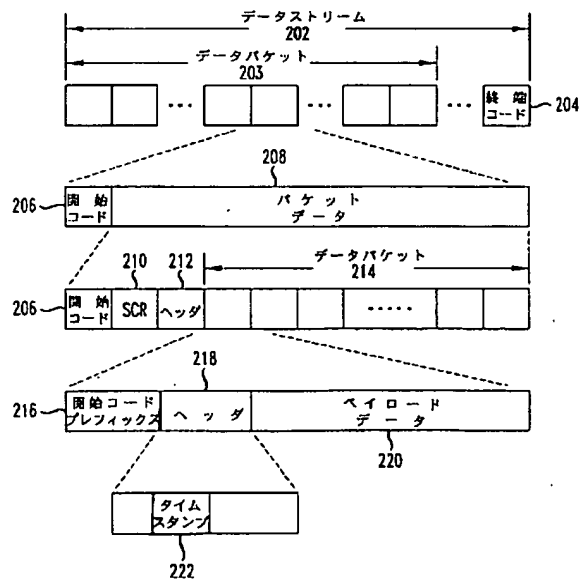
【図1】



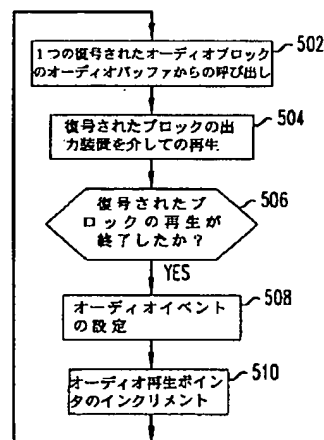
【図4】



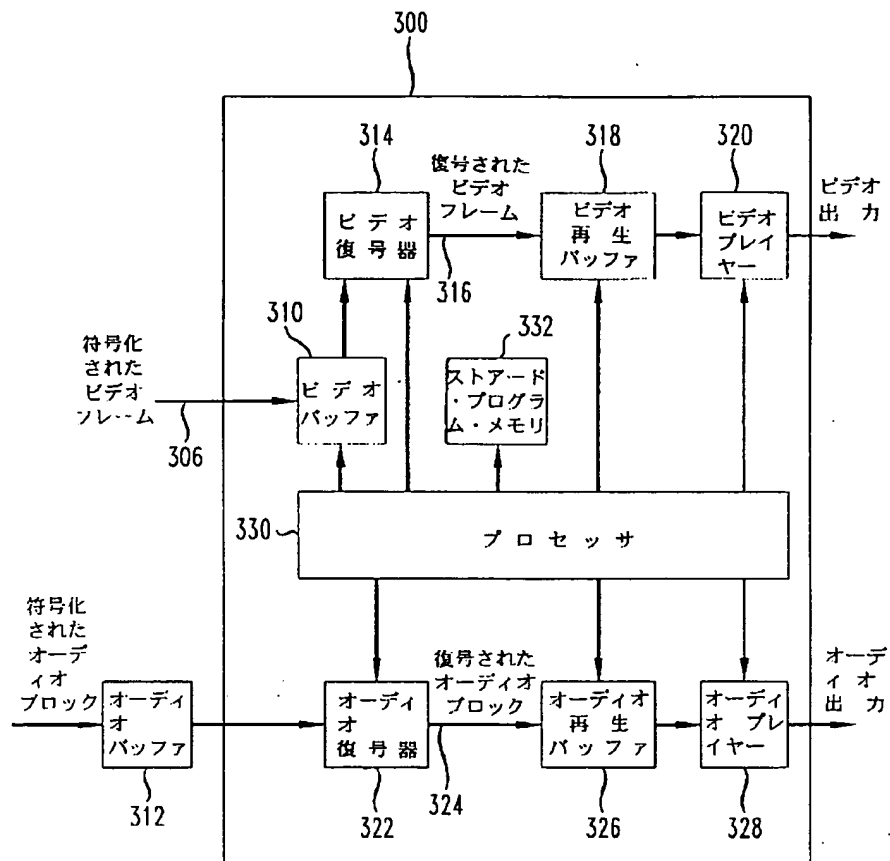
【図2】



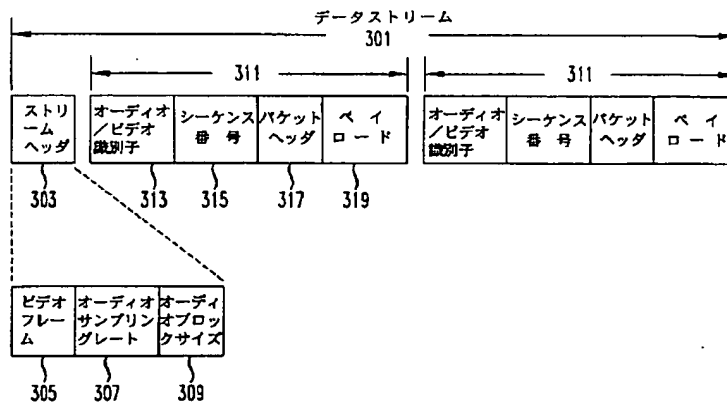
【図5】



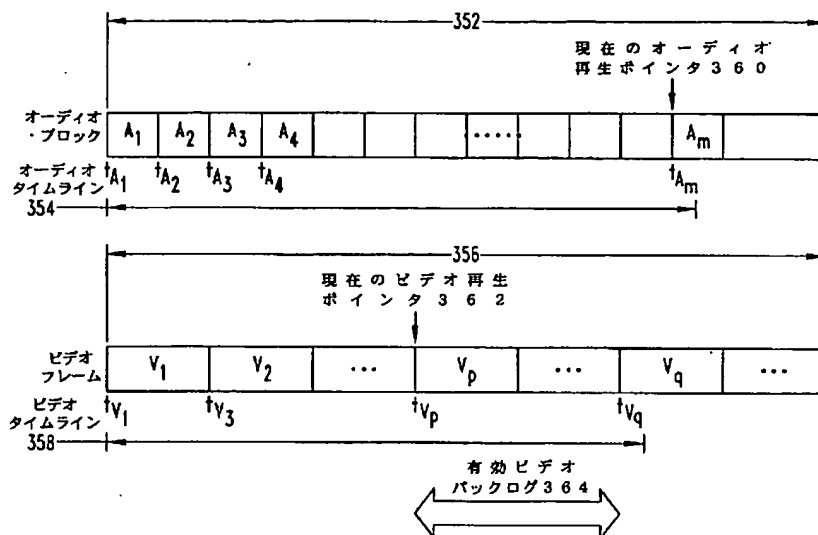
【図3A】



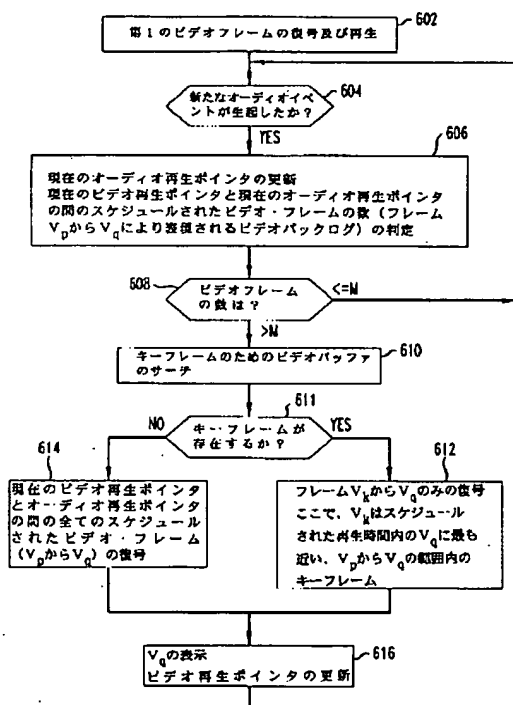
【図3B】



【図3C】



【図6】



フロントページの続き

(72)発明者 レイモンド ケイ. ジョンス
アメリカ合衆国 08043 ニュージャージー
ィ, ヴォーヒース, リーガン コート 15